

DIALOG(R) File 351:Derwent WPI
(c) 2000 Derwent Info Ltd. All rts. reserv.

011986562 **Image available**

WPI Acc No: 1998-403472/199835

XRPX Acc No: N98-314458

Image processing method for video conference system, video camera, video data editor - detects position of predetermined image in area limited by edge detector

Patent Assignee: CANON KK (CANO)

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 10162118	A	19980619	JP 96319891	A	19961129	199835 B

Priority Applications (No Type Date): JP 96319891 A 19961129

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 10162118	A	14		G06T-001/00	

Abstract (Basic): JP 10162118 A

The method involves performing band division of image data from an image feed zone (10), by a band division unit (20). An edge detector (30) detects the edge of the divided image data and limits the search area for a predetermined image.

A symmetrical axis detector (40) detects object property of the search area. The position of the predetermined image is detected in the area limited by the edge detector.

ADVANTAGE - Performs stable image detection. Reduces burden of search process.

BEST AVAILABLE COPY

This Page Blank (uspto)

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-162118

(43)公開日 平成10年(1998)6月19日

(51)Int.Cl.⁶

識別記号

F I

G 06 T 1/00

G 06 F 15/62

3 8 0

7/00

H 04 N 7/15

H 04 N 7/15

7/18

H

7/18

G 06 F 15/70

3 3 0 Q

審査請求 未請求 請求項の数13 OL (全 14 頁)

(21)出願番号

特願平8-319891

(71)出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(22)出願日

平成8年(1996)11月29日

(72)発明者 近藤 俊明

オーストラリア 2008 ニュー サウス
ウエールズ チッペンデイル シティーロ
ード 96 インターナショナル ハウス

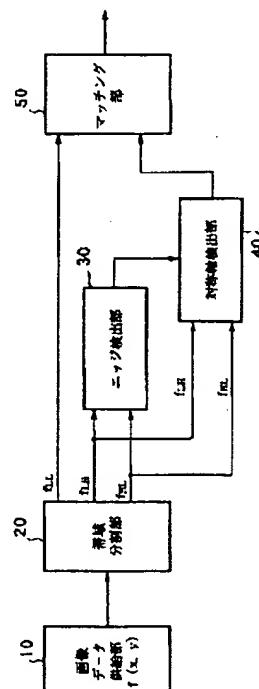
(74)代理人 弁理士 大塚 康徳 (外1名)

(54)【発明の名称】 画像処理装置及び方法

(57)【要約】

【課題】背景情報を用いずに、一枚の画像データからの特定像の検出を、簡易な構成でかつ高速に行うことを可能とする。

【解決手段】画像データ供給部10よりの画像データは帯域分割部20により、水平方向の高周波成分と垂直方向の低周波成分を含むf HLと、垂直方向の高周波成分と水平方向の低周波成分を含むf LH、低周波成分からなるf LLに分される。エッジ検出部30は、f LH及びf HLを用いてエッジがT字型に存在する部分を検出する。更に、対称軸検出部40では、エッジ検出部30で検出された部分について対象性の度合を算出し、所定の度合以上の対象性を有する部分が抽出される。マッチング部50は、以上のようにして限定された部分について顔の標準パターンとのマッチングを取り、顔の位置を検出する。



【特許請求の範囲】

【請求項1】 画像データから自動的に所定像を検出する画像処理方法であって、

前記画像データを帯域分割する分割工程と、

前記分割工程によって帯域分割された画像データに基づいて、前記所定像を探索すべき領域を限定する限定工程と、

前記限定工程で限定された領域において前記所定像の位置検出を行う検出工程とを備えることを特徴とする画像処理方法。

【請求項2】 前記所定像は、人間の顔であることを特徴とする請求項1に記載の画像処理方法。

【請求項3】 前記限定工程は、前記画像データより、エッジがT字型に分布した領域を選択的に抽出することを特徴とする請求項2に記載の画像処理方法。

【請求項4】 前記限定工程は、帯域分割された画像データに基づいて画像のエッジ部分を抽出し、該画像を所定個数の画素からなる複数のブロックに分割し、抽出された前記エッジ部分の画素を所定数以上含むブロックを前記所定像を探索すべき領域として決定することを特徴とする請求項2に記載の画像処理方法。

【請求項5】 前記限定工程は、前記エッジ部分の画素を所定数以上含むブロックがほぼT字状に連なる部分を抽出し、これを前記所定像を探索すべき領域とすることを特徴とする請求項4に記載の画像処理方法。

【請求項6】 前記限定工程は、当該画像の各部における左右の対称の度合に基づいて前記所定像を探索すべき領域を限定することを特徴とする請求項2に記載の画像処理方法。

【請求項7】 前記対象の度合は、画像データ値、画像データの勾配値、画像データの勾配の方向の少なくとも一つを用いて検出されることを特徴とする請求項6に記載の画像処理方法。

【請求項8】 前記限定工程は、前記画像データより、エッジがT字型に分布した領域を選択的に抽出し、該抽出された領域の各部について左右の対称の度合を検出し、該検出された度合に基づいて前記所定像を探索すべき領域を限定することを特徴とする請求項2に記載の画像処理方法。

【請求項9】 前記分割手段は、前記画像データを所定の空間周波数帯に分割することを特徴とする請求項2に記載の画像処理方法。

【請求項10】 前記分割手段は、前記画像データを：水平方向の高周波成分と垂直方向の低周波成分を含む第1の画像データと、垂直方向の高周波成分と水平方向の低周波成分を含む第2の画像データに分割することを特徴とする請求項8に記載の画像処理方法。

【請求項11】 前記分割手段は、更に、前記分割された画像データに間引き処理を行い、分割された画像データの画像サイズを減少することを特徴とする請求項8に

記載の画像処理方法。

【請求項12】 画像データから自動的に所定像を検出する画像処理装置であって、

前記画像データを帯域分割する分割手段と、

前記分割手段によって帯域分割された画像データに基づいて、前記所定像を探索すべき領域を限定する限定手段と、

前記限定手段で限定された領域において前記所定像の位置検出を行う検出手段とを備えることを特徴とする画像処理装置。

【請求項13】 画像データから自動的に所定像を検出するための制御プログラムを格納したコンピュータ可読メモリであって、前記制御プログラムがコンピュータを、

前記画像データを帯域分割する分割手段と、

前記分割手段によって帯域分割された画像データに基づいて、前記所定像を探索すべき領域を限定する限定手段と、

前記限定手段で限定された領域において前記所定像の位置検出を行う検出手段として機能させることを特徴とするコンピュータ可読メモリ、

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は画像データから人の顔等の所定像の位置を自動的に検出する画像処理装置及び方法に関するものである。更に詳しくは、本発明は、画像データから人の顔等の所定像の位置を自動的に検出し、人の顔の追尾機能や、伝送、圧縮する画像の特定、或いは焦点合わせや肌色調節の自動化等の機能で、監視カメラ、テレビ会議システム、ビデオフォン、ビデオカメラ、ビデオデータ編集装置等に適用可能な画像処理方法及び装置に関するものである。

【0002】

【従来の技術】従来、顔の自動検出方法及び装置では、以下に示す手法が取られている。すなわち、

①既知の背景との差分を求め、この差分値の大きい位置に顔が存在するものとして顔の自動検出を行う、

②テレビ映像のフレームあるいはフィールド間差を求め、差の大きい部分に顔が存在するものとして顔の自動検出を行う、

③カラー映像信号から肌色の領域を探すことでの顔の自動検出を行う、

④パターンとして何らかの顔に関する何らかの知識（例えば、輪郭モデルや統計的特徴）に基づいて類似パターンを探索することで顔の自動検出を行う、

⑤無地な背景を前提として、画像データの水平、垂直方向への投影値を求めて、この投影値に基づいて顔の自動検出を行う、といった手法が一般的である。

【0003】

【発明が解決しようとする課題】しかしながら、上記従

米例の①～⑤の手法では、次に述べるような課題がある。すなわち、①の手法では、背景情報を既知としているため、応用範囲がテレビ会議や監視カメラ等、所定の場所に設置されたものに限られてしまう、②の手法では、映像信号の時間差を求めるために高価な画像メモリが必要となる。また、人以外の動物体や移動体が画面に含まれたり、映像入力装置（例えばカメラ）本体が動くような場合は、正しい検出が困難となる、③の手法では、照明条件によって肌色も多様に変化するので、照明に厳しい条件が課せられることになる、④の手法では、顔パターンと類似した領域を例えば相関演算により画像の広い範囲で探索することが必要となる。このため演算量が多く、低速であったり大規模な装置が必要となってしまう、⑤の手法では背景が無地の場合に応用が限られてしまい、汎用性に乏しい、といった問題がある。

【0004】本発明は、上記問題に鑑みてなされたものであり、背景情報を用いずに、一枚の画像データからの特定像の検出を、簡易な構成でかつ高速に行なうことを可能とする画像処理方法及び装置を提供することを目的とする。

【0005】また、本発明の他の目的は、照明条件の変化による影響を受けにくく、安定した像検出を行える画像処理方法及び装置を提供することを目的とする。

【0006】また、本発明の他の目的は、簡易なエッジ検出、及びエッジ処理により素早く画像情報の乏しい無地の領域、低コントラストな領域を顔探索の対象領域から除去することを可能とし、高速な顔検出を可能とする画像処理方法及び装置を提供することにある。

【0007】また、本発明の他の目的は、複雑な背景から左右対称の領域を抽出することにより、後段の顔探索の負担を大幅に減ずることを可能とし、高速な顔検出を可能とする画像処理方法及び装置を提供することにある。

【0008】

【課題を解決するための手段】上記目的を達成するための本発明の画像処理方法は以下の工程を備える。すなわち、画像データから自動的に所定像を検出する画像処理方法であって、前記画像データを帯域分割する分割工程と、前記分割工程によって帯域分割された画像データに基づいて、前記所定像を探査すべき領域を限定する限定工程と、前記限定工程で限定された領域において前記所定像の位置検出を行う検出工程とを備える。

【0009】また、上記目的を達成する本発明の画像処理装置は、画像データから自動的に所定像を検出する画像処理装置であって、前記画像データを帯域分割する分割手段と、前記分割手段によって帯域分割された画像データに基づいて、前記所定像を探査すべき領域を限定する限定手段と、前記限定手段で限定された領域において前記所定像の位置検出を行う検出手段とを備える。

【0010】

【発明の実施の形態】以下に説明する実施形態によれば、帯域分割手段が処理対象の画像データを、水平・垂直方向の高周波成分、低周波成分へ分割する。そして、顔領域限定手段が、当該帯域分割手段によって分割された画像データを利用して、顔の存在する可能性のある領域（顔領域）を限定していく。ここで、帯域分割手段においては、画像の解像度もしくはサンプリングブレートを落とす処理が行われる。そして、顔領域限定手段が当該帯域分割手段の出力信号を直接利用して顔領域を限定する。このため、高速で簡便な顔の検出方法及び装置を提供できる。

【0011】また、顔領域限定手段により、エッジがT字型に分布する部分が抽出される。このため、エッジ数の乏しい低コントラストな領域とT字型にエッジが分布しない無地な背景上の物体の輪郭域が顔領域から除去される。よって、比較的平坦な背景を持った画像からの顔検出を容易に行える。また、エッジ検出には、上記帯域分割手段の出力のうちの、水平方向の高周波成分及び垂直方向の高周波成分が直接利用されるので極めて高速な処理ができる。

【0012】更に、顔領域限定手段によれば、左・右対称の度合の強い領域が抽出される。複雑な背景を持った画像でも、左・右対称という特異な領域を抽出することにより、顔の存在領域を効果的に限定することができる。また、対称形探索には、上記帯域分割手段の出力を直接利用するので、極めて効率的な処理が可能である。

【0013】以下、添付の図面を参照して本発明の好適な実施形態をより具体的に説明する。

【0014】（第1の実施形態）図1は本実施形態の画像処理装置の概略の構成を示すブロック図である。同図において、101はCPUであり、ROM102等に格納された制御プログラムを実行することにより、本画像処理装置の各種処理を実現する。102はROMであり、CPU101が実行する制御プログラムや各種データを格納する。103はRAMであり、CPU101が各種制御を実行する際の作業領域を提供する。104は外部記憶装置であり、例えばハードディスク、フロッピーディスク等で構成される。外部記憶装置104には、アプリケーションプログラムや画像データ等を格納することができる。105は入力部であり、キーボードやマウス等を備える。106はディスプレイであり、CPU101の制御により各種表示を行う。

【0015】107はカメラインターフェースであり、カメラ108より取り込まれた映像データをデジタルデータに変換してバス109へ出力する。108はカメラであり、CCD等の撮像素子を備える。109はバスであり、上述の各構成を接続する。なお、上記説明において、CPU101はROM102に格納されたプログラムを実行するとしたが、外部記憶装置104に格納されたプログラムをRAM103にロードして、これを

CPU104によって実行するように構成してもよい。また、カメラ108によって得られた画像データは外部記憶装置104或いはRAM103に格納される。また、画像を入力するための装置としては、上述のカメラ108に限らず、例えばスキャナ等であってもよい。110は画像データから顔の位置を検出する顔位置検出部である。

【0016】図2は第1の実施形態の画像処理装置における顔位置検出部に関する機能構成を示すブロック図である。同図において、10は画像データ供給部であり、当該システムへ画像データを供給する。なお、画像データ供給部10は、カメラ108(或いはスキャナ等)によって構成される画像データ入力部であってもよいし、画像データを格納する外部記憶装置104、RAM103によって構成される画像データ格納部であってもよい。20は帯域分割部、30はエッジ検出部、40は対称軸検出部、50は顔の位置を決定するマッチング部である。各部の機能の詳細は後述する。マッチング部50の出力は、顔の位置を示すデータとして出力される。

【0017】次に、図3を参照し、図2の機能構成の動作を信号の流れに沿って詳しく説明する。図3は第1の実施形態による顔位置検出の手順を示すフローチャートである。画像入力部あるいは画像データ格納部等の画像データ供給部10から画像データが入力される(ステップS10)。なお、帯域分割部20は、画像を各空間周波数帯に分割し、更にダウンサンプリング(間引き)により、画像サイズを縮小する。これは、ダウンサンプリングにより処理すべき画像データの量を減少させ、以後の処理をより高速にするためである。

【0018】以後、簡単のため、入力画像を $f(x,y)$ 、ダウンサンプリングされたサブ画像を f_{LL} 、 f_{LH} 、 f_{HL} 、 f_{HH} と表記する。ここで、 f_{LL} は $f(x,y)$ の水平垂直方向の低周波成分から成るサブ画像、 f_{LH} は $f(x,y)$ の水平方向の低周波成分と垂直方向の高周波成分から成るサブ画像、 f_{HL} は $f(x,y)$ の水平垂直方向の高周波成分と垂直方向の低周波成分から成るサブ画像、 f_{HH} は $f(x,y)$ の水平垂直方向の高周波成分から成るサブ画像である。図4はサブ画像の1例を示す図である。図4で左上が f_{LL} 、右上が f_{LH} 、左下が f_{HL} 、右下が f_{HH} にそれぞれ対応する。 f_{LL} 、 f_{HL} は共に、エッジ検出部30と対称軸検出部40へと入力される。また、 f_{LL} は顔の位置を決定するマッチング部50へと入力される。なお、 f_{HH} は画像情報が乏しいので本実施形態では使用しない。

【0019】さて、以上のようにして得られた f_{HL} 及び f_{LH} はエッジ検出部30に供給されて、エッジ検出が行われ、処理を行なべき領域として有効ブロックが抽出される(ステップS11)。図5はエッジ検出部30内の処理の概要を示すフローチャートである。また、図6はエッジ検出部30による処理のより詳細な手順を示すフローチャートである。70はエッジ検出処理であり、

入力されたサブ画像 f_{HL} 、 f_{LH} よりエッジ画像を獲得する。図5のフローチャートに従って説明すると、まず、サブ画像 f_{HL} 、 f_{LH} を入力し(ステップS701)、これらサブ画像に対して閾値処理を施して2値化を行い、 f'_{HL} 及び f'_{LH} を得る(ステップS702)。続いて、ステップS703において、以上のようにして2値化された f'_{HL} 、 f'_{LH} の論理和(OR)を演算する。こうして得られた画像がエッジ画像である。

【0020】図7はエッジ検出処理70の処理例を具体的に示す図である。左上のサブ画像は f_{LL} 、右上のサブ画像は f_{LH} を所定の閾値で2値化して得られたサブ画像 f'_{LH} である。また、右下のサブ画像は f_{LL} を所定の閾値で2値化して得られたサブ画像 f'_{LL} 、そして右下がこれら2つのサブ画像(f'_{LH} 及び f'_{LL})の論理和、即ち、エッジ画像を表している。

【0021】続いて、エッジ検出処理70によって得られたエッジ画像は、エッジブロック抽出処理80に供給される。エッジブロック抽出処理80では、前記エッジ画像を小さいブロック(例えば、1ブロック当たり 8×8 画素)に細かく分割し(ステップS801)、各ブロックに含まれるエッジ数をカウントする(ステップS802)。そして、カウントされたエッジ数に対して閾値処理を施すことにより、一定数以上のエッジを含むブロックを抽出する(ステップS803)。こうして抽出されたブロックをエッジブロックと呼ぶ。

【0022】図8はエッジ画像とエッジブロックの対応を説明する図である。図8の(a)は図7で示したエッジ画像であり、図8の(b)は図8の(a)のエッジ画像からエッジ部の画素を所定個数以上含むブロックを抽出した結果である。なお、本例では、ブロックサイズは 8×8 画素で、1ブロックの総計64画素中8画素以上がエッジである場合に、当該ブロックをエッジブロックとして黒く表示させている。図8より明らかのように、この時点では既に無地の背景は除外されることになる。

【0023】次に、顔エッジ抽出処理90では、前記処理80で抽出されたエッジブロックからT字型に分布したエッジブロックを抽出する(ステップS901)。顔エッジ抽出処理90内の処理を図9、図10を用いて更に詳しく説明する。

【0024】図9はT字型に分布したエッジブロックを表わす図である。顔から検出されるエッジはT字型に分布するので、このようなT字型に分布しているエッジブロックのみを選択的に抽出し、顔の検出対象領域を制限する。以下、このようにして抽出されたエッジブロックを有効エッジブロックという。こうして抽出された有効エッジブロックは、対称軸検出部40へ出力される。

【0025】図10は顔エッジ抽出処理90の出力例を示す図である。T字型にエッジブロックが分布する領域のみを選択することにより、孤立したブロック、平坦な背景を持つ物体の輪郭部に位置するブロックが除去さ

れ、以後の処理（対称軸検出部40による処理、マッチング部50による処理）で、処理対象とすべき領域を大幅に制限することができる。

【0026】なお、本例では注目とするブロックを中心には左・右・下方向に2個ずつのブロックから成るT字型（図9）を用いたが、当設定は予想される顔のサイズ、用いるブロックのサイズに応じて変更されるものである。また、注目するT字型の近隣ブロック（本例の場合、注目ブロックと6個のブロック）全てがエッジブロックである必要はない。例えば、注目ブロックと注目すべき近隣ブロックの内、一定数以上がエッジブロックであるならば有効エッジブロックであるとして、対応するエッジブロックを選択するというように、緩い基準設定をしても良い。なお、これらの基準設定は、入力部10より行う。

【0027】更に、入力する顔のサイズが全く不明の場合に備え、例えば画面サイズから 12×12 画素の最少サイズまでの間で何段階かの異なるサイズで、エッジ検出処理70、エッジブロック抽出処理80、顔エッジ抽出処理90を実行することも可能である。このように、複数段階のサイズで各々の処理を行い、各サイズの処理で求められる対称性の度合と相関係数 ρ の値が高いサイズが、その時の入力画像に適切なサイズとして選択され、その後の処理に用いられる。以上のようなエッジ検出部30の出力は、引続く処理で対象とする領域を示すマップとして、対称軸検出部40の演算領域を制御する。

【0028】さて、ステップS12（図3）では、以上のようにして抽出された有効ブロックの領域の各部において対称性の度合を検出する。以下、対称軸検出部40の処理を説明する。図11は対称軸検出部40の処理手順を説明するフローチャートである。対称軸検出部40では、上記エッジ検出部30によって有効とされた領域でのみ実行される。

【0029】まず、帯域分離部20より f_{HL} 及び f_{LH} を入力するとともに、エッジ検出部30より有効エッジブロックを入力する（ステップS401）。そして、これ

$$\begin{aligned} r(x, y) = & \sum_{j=y-\frac{v}{2}}^{y+\frac{v}{2}} \sum_{i=x-\frac{h}{2}}^x \left\{ \left| \cos\{\theta(i, j)\} + \cos\{\theta(2x-i, j)\} \right|^2 \right. \\ & \left. + \left| \sin\{\theta(i, j)\} - \sin\{\theta(2x-i, j)\} \right|^2 \right\} \quad \cdots (4) \end{aligned}$$

【0038】ここで、(4)式は当ブロック内の左・右のθ差を演算する。なお、演算ブロック範囲は検出すべき顔のサイズとほぼ一致させる。顔のサイズが未知の場合には何段階か異なるサイズのBを用いても良い。図12は帯域分離された画像データ f_{HL} (a)と対称性の度合 r (b)との対応関係を示す図である。対称性が高い場合、(4)式の第1項、第2項は共に、項内の相殺に

より小さな値を取る。また、(4)式は、絶対値を更に2乗したものであるが、対称性の度合 r は、単に絶対値和であってもよい。

【0039】図12の(b)の暗い垂直方向のストライプが顔の対称軸を示している。図12の(b)では説明のため画像全域で r を演算し表示しているが、実際には対象性の度合を求めるための演算はエッジ検出部30で

$$\theta(x, y) = \tan^{-1} \left(\frac{\partial f(x, y) / \partial y}{\partial f(x, y) / \partial x} \right) \quad \cdots (1)$$

【0031】ここで、上記(1)式の右辺の勾配は、画像の高周波成分と等価であり、以下の(2)式が成立する。

【0032】

【数2】

$$f_{LH} = \frac{\partial f(x, y)}{\partial y}, \quad f_{HL} = \frac{\partial f(x, y)}{\partial x} \quad \cdots (2)$$

【0033】従って、本実施形態では、勾配ベクトルの方向 θ を、帯域分割部20の出力 f_{LH} 、 f_{HL} を直接利用して、(1)式の替わりに、以下の(3)式を実行して求める。

【0034】

【数3】

$$\theta(x, y) = \tan^{-1} \left(\frac{f_{LH}}{f_{HL}} \right) \quad \cdots (3)$$

【0035】上記の(3)式は、ソフトウェアで実行しても良いし、ルックアップテーブルを利用してハードウェアで実行しても良い。

【0036】次に、ステップS403において、画像の各位置 (x, y) における対称性の度合 r を求める。 r (x, y) は、 $\theta(x, y)$ を用いて次の(4)式のように定義される。

【0037】

【数4】

有効エッジブロックが抽出された領域でのみ実行される。図13は有効エッジブロックにおいて対象性の度合 γ を算出した状態を示す図である。すなわち、図13は図10と図12の論理積に等しくなる。

【0040】さて、以上のようにして対象性の度合 γ を算出すると、ステップS404(図10)に進み、算出された γ が所定の閾値以上の値を有する部分を抽出する。図13は γ に閾値処理を施し、顔である可能性の高い領域(顔候補領域)をより限定した結果を示す図である。図14中の黒い島の部分において、マッチング部50(図2)による最終的な処理が実行され、顔の位置が

$$\sigma(x, y) = \frac{\sum_j \sum_i [f(i, j) - \bar{f}] [w(i - x, j - y) - \bar{w}]}{\sqrt{\sum_j \sum_i [f(i, j) - \bar{f}]^2 \sum_j \sum_i [w(i - x, j - y) - \bar{w}]^2}} \quad \dots (5)$$

【0044】但し、 $f(x, y)$ は画像、 $w(x, y)$ はテンプレート、 f バーは $\omega(x, y)$ と重なっている $f(x, y)$ の平均値、 w バーは $w(x, y)$ の平均値である。

【0045】図14は図14の領域の内、上記正規化相互相関係数 σ が0.35以上である領域を示す図である(○は(5)式により-1~+1にスケーリングされており、1に近い程、類似)。

【0046】図15に示される黒い島は、上記(5)式によって処理を行った結果、顔である可能性が極めて高い領域を示している。そして、図16の白い四角は σ が最大の場所を前記fLL上に重ねて表示させた図である。ベストマッチング点が顔の位置(両目の間)を正しく検出していることがわかる。

【0047】マッチング部50による顔の判定処理は、上述したコリレーション(相関演算)の他に顔パターンの多少の変形を許容する弹性マッチング、フーリエ変換法等、既存のいかなる方法であっても良く、特に上記実施形態に限定されるものではない。

【0048】以上説明したように、上記実施形態によれば、有効エッジブロックを抽出することで対象性の検出処理の処理量を低減し、更に、対象性の度合の高い部分のみにパターンマッチングを施して顔の位置を決定することで、パターンマッチングの処理量を大幅に低減している。このため、高速な顔位置検出が可能となる。

【0049】更に、第1の実施形態によれば、
1. 帯域分割部20(図2)の活用によりサンプリングレートを落としているので、顔探索を効率的に実行できる。
2. 帯域分割部20の出力を、後段のエッジ検出部30、対称軸検出部40(図2)で直接的に利用するので、演算が高速でハードウェア化も容易となる。特に無地の背景、低コントラスト領域除去に効果がある。
3. 対称軸検出部40は、画像データの方向を利用していているので、照明条件の変化に影響を受けにくく、

決定される(ステップS13、S14)。

【0041】マッチング部50は、エッジ検出部30、対称軸検出部40による処理の結果、限定された顔候補領域内(図14)で最も顔らしい位置を決定する。具体的には典型的な顔をテンプレートとして保持しておく、当該テンプレートと最も類似した場所を探索する。

【0042】類似度判定の基準は照明条件の変化に強い正規化相互相関係数を用いる。当係数 $\sigma(x, y)$ は次の(5)式で定義される。

【0043】

【数5】

安定した対称軸検出ができる。また、左・右対称形という特異な領域を探査するので、複雑な背景であっても顔領域を大幅に限定できるという効果がある。

【0050】また、帯域分割部20による処理を繰り返し適用することにより、一つの入力画像に対する解像度の異なる画像を生成することができる。これらの解像度の異なる画像各々にエッジ検出部30及び対称軸検出部40の処理を実行し、この中から、適切な情報を選択することにより、先に述べた、複数段階のブロックサイズを用いてスケーリングの異なる画像が入力される場合に対処するとの同様の効果が得られる。

【0051】(第2の実施形態)上記第1の実施形態では、顔領域を限定する手法として、エッジ検出部30(図2)と対称軸検出部40(図2)を設けた。しかしながら、背景が無地であるような環境下では、エッジ検出部30単独で顔領域の限定を行うこともできる。また、逆に複雑な背景を対象とする場合には、対称軸検出部単独で顔領域の限定を行っても良い。このように、第2の実施形態によれば、取り扱う背景環境に応じて、不必要的処理を取り除くことが可能となり、より効率的で高速な顔検出が実現できる。

【0052】図17は第2の実施形態の画像処理装置における顔位置検出部に関する機能構成を示すブロック図である。同図において第1の実施形態(図2)と同様の構成には同じ参照番号を付し、ここでは説明を省略する。第2の実施形態では、顔の位置検出に際して、モード1、2、3の3通りを設定できる。モード1は第1の実施形態と同様に、エッジ検出部30によって抽出された有効ブロックの領域に対して対称軸検出部40が対称性の度合 γ を検出し、所定値以上の γ 値を有する領域についてマッチング部50がマッチング処理を実行する。モード2では、エッジ検出部30によって検出された有効ブロックの領域についてマッチング部50によるマッチング処理を実行する。また、モード3では、対称軸検出部30によって抽出された有効ブロックの領域に対して対称軸検出部40が対称性の度合 γ を検出し、所定値以上の γ 値を有する領域についてマッチング部50がマッチング処理を実行する。

部40が入力画像の全領域にわたって対称性の度合 γ を算出し、マッチング部50が所定値以上の γ 値を有する部分にマッチング処理を実行する。

【0053】図17において、100は切換器であり、対称軸検出部40が対称軸の検出（対称性の度合の検出）を行う領域を指定する信号を切り換える。モード1では、エッジ検出部30よりの有効ブロックの領域について対称軸検出部40が処理を行うので、エッジ検出部30の出力を対称軸検出部40に供給するべくスイッチはa側に切り換える。また、モード3では対称軸検出部40が全領域について対称性の度合 γ を算出するので、切換器100のスイッチはb側に切り換える。

【0054】切換器110もモードに応じて切り換えが制御される。すなわち、モード2では、エッジ検出部30によって抽出された有効ブロックの領域についてマッチング部50が処理を施すので切換器110のスイッチはc側に切り換える。一方、モード1、3においては、対称軸検出部40によって対称性の度合 γ が所定値以上となった部分を処理するので、切換器110のスイッチはd側に切り換えられる。

【0055】図18は第2の実施形態の処理の流れを説明するフローチャートである。同図において、図3と同じ処理ステップには同一のステップ番号を付してある。同図に示されるように、モード1が指定された場合（モードの指定は入力部105より行うものとする）は、切換器100がa側に、切換器110がd側に切り換えられる。この結果、ステップS21、S22の分岐を経由して図3の手順と同様の処理が実行されることになる。モード2の場合は、切換器110がc側に切り換えられるので、ステップS22よりステップS23へ処理が進み、マッチング部50がエッジ検出部30で抽出された有効ブロックの領域についてマッチング処理を実行する。そして、ステップS14で、マッチング度のもっとも高い部分（ベストマッチング点）を顔の位置として決

定する。更に、モード3が指定された場合は、切換器100がb側に、切換器110がd側に切り換えられる。このため、処理はステップS21からステップS24へ進み、画像全域についてf_{LH}及びf_{HL}から対称性の度合 γ が算出される。そして、ステップS13では、マッチング部50が、度合 γ が所定値以上の部分についてマッチングを行い、ステップS14でベストマッチング点を顔の位置として決定する。

【0056】以上のように、第2の実施形態によれば、顔位置検出のためのマッチング処理の実行範囲を限定する処理の内容を背景等の状態に応じて設定できる。このため、より効率的で高速な顔位置検出処理が実現する。

【0057】なお、上記各実施形態で説明した顔位置検出処理は、専用のハードウェアで実現してもよいし、上述の制御を実現するためのソフトウェアをCPU101が実行することで実現されてもよいし、ソフトウェアとハードウェアを適切に混在させることで実現されてもよい。なお、ソフトウェアはROM102に格納されてもよいし、外部記憶装置104よりRAM103にロードされてもよい。

【0058】（第3の実施形態）第1の実施形態では対称軸検出部40（図2）の対称性の指標として、画像勾配の方向性を用いた。しかしながら、対称軸検出の方法は、このような方法に限られるものではない。例えば、勾配の絶対値の左・右差を求める（6）式及び勾配値そのものの左・右差を求める（7）式により画像勾配値そのものを直接用いて、ブロック内の左・右差を演算しても良いし、より簡単には、左・右差の絶対値和を求める（8）式或いは（9）式、また、或いは左右のクロスコリレーション（相関演算）の（10）式により、画像値（輝度値）を用いて対称性を演算しても良い。

【0059】

【数6】

$$\gamma(x, y) = \sum_{j=y-\frac{v}{2}}^{y+\frac{v}{2}} \sum_{i=x-\frac{h}{2}}^x \left\{ |f_{HL}(i, j)| - |f_{HL}(2x-i, j)| + |f_{LU}(i, j)| - |f_{LU}(2x-i, j)| \right\} \quad \dots(6)$$

【0060】

【数7】

$$\gamma(x, y) = \sum_{j=y-\frac{v}{2}}^{y+\frac{v}{2}} \sum_{i=x-\frac{h}{2}}^x \left\{ |f_{HL}(i, j) + f_{HL}(2x-i, j)| + |f_{LU}(i, j) - f_{LU}(2x-i, j)| \right\} \quad \dots(7)$$

【0061】

【数8】

$$\gamma(x, y) = \sum_{j=y-\frac{v}{2}}^{y+\frac{v}{2}} \sum_{i=x-\frac{h}{2}}^x |f_{LU}(i, j) - f_{LU}(2x-i, j)| \quad \dots(8)$$

【0062】

$$\gamma(x, y) = \sum_{j=y-\frac{v}{2}}^{y+\frac{v}{2}} \sum_{i=x-\frac{h}{2}}^{x+\frac{h}{2}} [f_{LU}(i, j) - f_{LU}(2x-i, j)]^2 \quad \cdots (9)$$

【0063】

$$\gamma(x, y) = \frac{\sum_{j=y-\frac{v}{2}}^{y+\frac{v}{2}} \sum_{i=x-\frac{h}{2}}^{x+\frac{h}{2}} [f_{LU}(i, j) - \bar{f}_L][f_{LU}(i, j) - \bar{f}_R]}{\sqrt{\sum_{j=y-\frac{v}{2}}^{y+\frac{v}{2}} \sum_{i=x-\frac{h}{2}}^{x+\frac{h}{2}} [f_{LU}(i, j) - \bar{f}_L]^2 [f_{LU}(2x-i, j) - \bar{f}_R]^2}} \quad \cdots (10)$$

ただし、 \bar{f}_L, \bar{f}_R は、それぞれブロック($h \times v$)内の
左半分、右半分の平均値とする。

【0064】以上のように第3の実施形態によれば、画像勾配値を直接用いることにより、当該勾配値の方向 θ を演算するプロセスを省くことができ、より簡潔な回路構成（或いはプログラム構成）とすることが可能となる。画像値を用いる方法では、例えば前記ブロック内の左・右間でクロスコリレーションを行うとすれば、相関演算に必要なブロック内の平均値と分散値を最終段のマッチング部50（図2）と共有することができる。従って、計算量の多い相関演算量の負担を減らす効果がある。

【0065】なお、本発明は、複数の機器（例えばホストコンピュータ、インタフェイス機器、リーダ、プリンタなど）から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

【0066】また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0067】この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0068】プログラムコードを供給するための記憶媒体としては、例えば、フロッピディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

【0069】また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示

【数9】

【数10】

に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0070】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0071】本発明を上記記憶媒体に適用する場合、その記憶媒体には、先に説明したフローチャートに対応するプログラムコードを格納することになるが、簡単に説明すると、図19のメモリマップ例に示す各モジュールを記憶媒体に格納することになる。

【0072】すなわち、少なくとも「分割処理モジュール」「限定処理モジュール」および「検出処理モジュール」の各モジュールのプログラムコードを記憶媒体に格納すればよい。ここで、コレラモジュールは、画像データから自動的に顔画像の位置を検出すための制御プログラムを構成するものであり、分割処理モジュールは当該画像データを帯域分割する分割処理を実現する。また、限定処理モジュールは、分割処理によって帯域分割された画像データを用いて顔画像を探索すべき領域を限定する限定処理を実現する。そして、検出処理モジュールは、限定処理で限定された領域において顔の位置検出を行う検出処理を実現する。

【0073】以上説明したように、上記各実施形態によれば、高速で効率的な顔の検出をすることができる。また、エッジ検出部30の処理により、低コントラストな領域や無地領域を高速に除去することができる。更に、対称軸検出部40を設けたことにより、複雑な背景の中

でも、左・右対称形という特異な領域を抽出することができ、顔の存在領域を大幅に限定することができる。特に、上記実施形態によれば、対称軸検出部40による対称軸の検出が画像データの勾配の方向を利用しているので、照明条件の変化に影響を受けにくく、左・右対称領域をより確実に抽出することができる。

【0074】また、本発明を監視カメラやテレビ会議システム、ビデオフォンに適用することにより、自動追尾機能を実現させたり、或いは所定画像のみの伝送、圧縮を可能とし、伝送能力を向上させることができる。

【0075】また、本発明をビデオカメラに適用することにより、所定像への自動焦点合わせや肌色自動調節を実現できる。

【0076】また、本発明をビデオ編集装置に適用することで、膨大なデータの中から人の顔等の所定像だけを自動的に検出でき、編集作業の大幅な短縮が可能となる。

【0077】更に、本発明は、セキュリティが求められるビル、オフィス、ATMや端末への不法なアクセスの阻止、税関でのパスポート写真と本人の照合、顔のデータベースの検索等の自動化を可能とし、マンマシンインターフェースへの応用も可能である。

【0078】

【発明の効果】以上説明したように本発明によれば、背景情報を用いずに、一枚の画像データからの所定像の検出を、簡易な構成でかつ高速に行うことが可能となる。

【0079】また、本発明によれば、照明条件の変化による影響を受けにくく、安定した像検出が行える。

【0080】また、本発明によれば、簡易なエッジ検出、及びエッジ処理により素早く画像情報の乏しい無地の領域、低コントラストな領域を顔探索の対象領域から除去することができる。このため、顔検出の対象となる領域が限定され、高速な顔検出が可能となる。

【0081】また、本発明によれば、複雑な背景から左右対称の領域を抽出することにより、後段の顔探索処理の負担を大幅に減ずることが可能となり、高速な顔検出が実現される。

【0082】

【図面の簡単な説明】

【図1】本実施形態の画像処理装置の概略の構成を示すブロック図である。

【図2】第1の実施形態の画像処理装置における顔位置検出部に関する機能構成を示すブロック図である。

【図3】第1の実施形態による顔位置検出の手順を示すフローチャートである。

【図4】サブ画像の1例を示す図である。

【図5】エッジ検出部30内の処理の概要を示すフローチャートである。

【図6】エッジ検出部30による処理のより詳細な手順を示すフローチャートである。

【図7】エッジ検出処理70の処理例を具体的に示す図である。

【図8】エッジ画像とエッジブロックの対応を説明する図である。

【図9】丁字型に分布したエッジブロックを表わす図である。

【図10】顔エッジ抽出処理90の出力例を示す図である。

【図11】対称軸検出部40の処理手順を説明するフローチャートである。

【図12】帯域分離された画像データfLLと対称性の度合 α との対応関係を示す図である。

【図13】 α に閾値処理を施し、顔である可能性の高い領域（顔候補領域）をより限定した結果を示す図である。

【図14】図13に示した状態より、対象性の度合 α が所定値以上の部分を抽出した状態を示す図である。

【図15】図14に示した部分より、顔の標準パターンとのマッチング度の高い部分を抽出した状態を示す図である。

【図16】図15に示した部分において、最もマッチング度の高い部分と元画像との対応を示す図である。

【図17】第2の実施形態の画像処理装置における顔位置検出部に関する機能構成を示すブロック図である。

【図18】第2の実施形態の処理の流れを説明するフローチャートである。

【図19】本発明に係る制御プログラムを格納した記憶媒体のメモリマップ例を示す図である。

【符号の説明】

10 画像データ供給部

20 帯域分割部

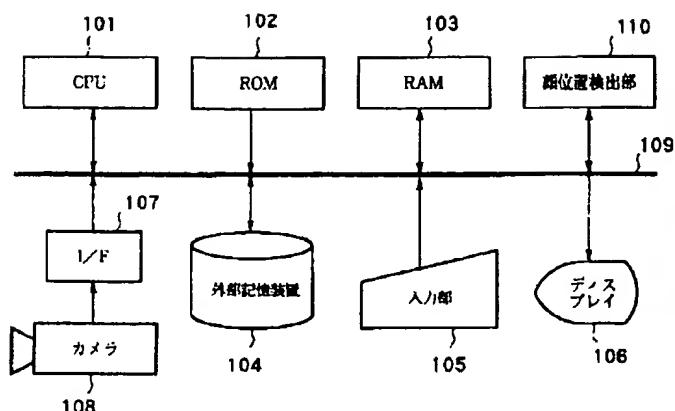
30 エッジ検出部

40 対称軸検出部

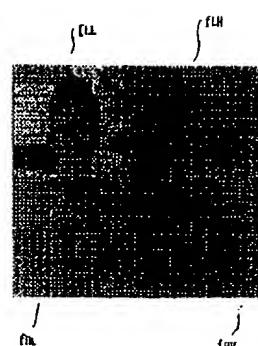
50 マッチング部

60 出力端子

【図1】

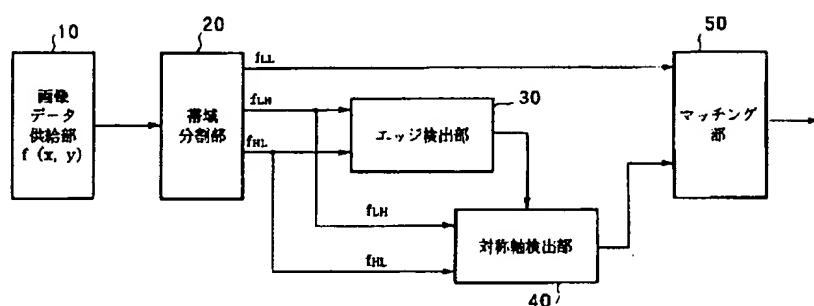


【図4】



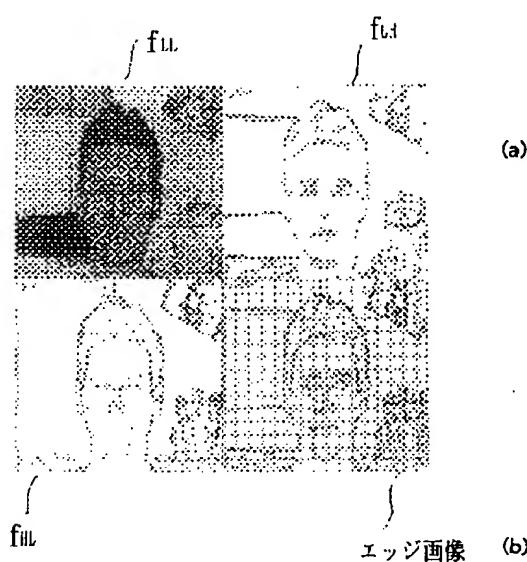
【図9】

【図2】

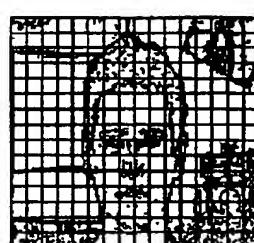


【図10】

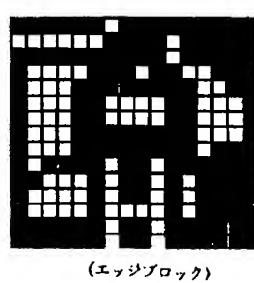
【図7】



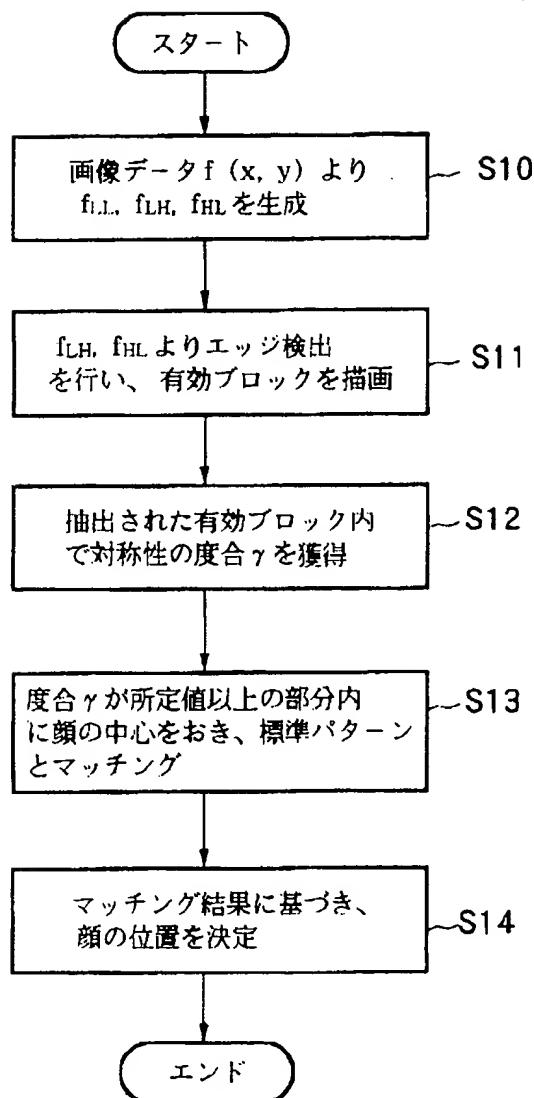
【図8】



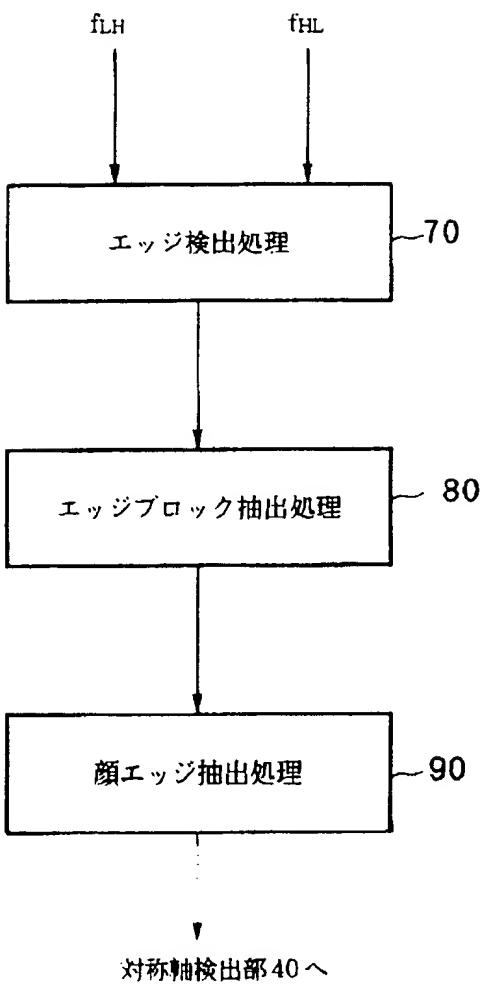
【図13】



【図3】

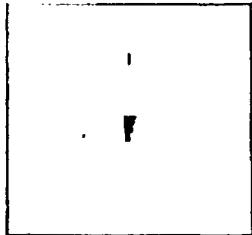


【図5】

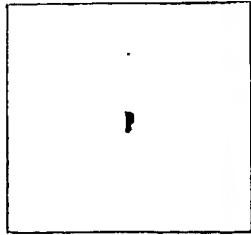


【図19】

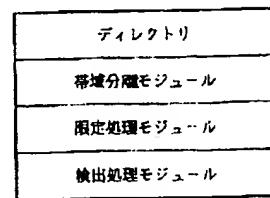
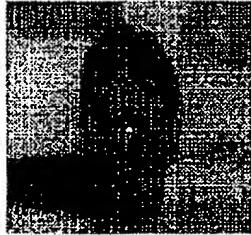
【図14】



【図15】

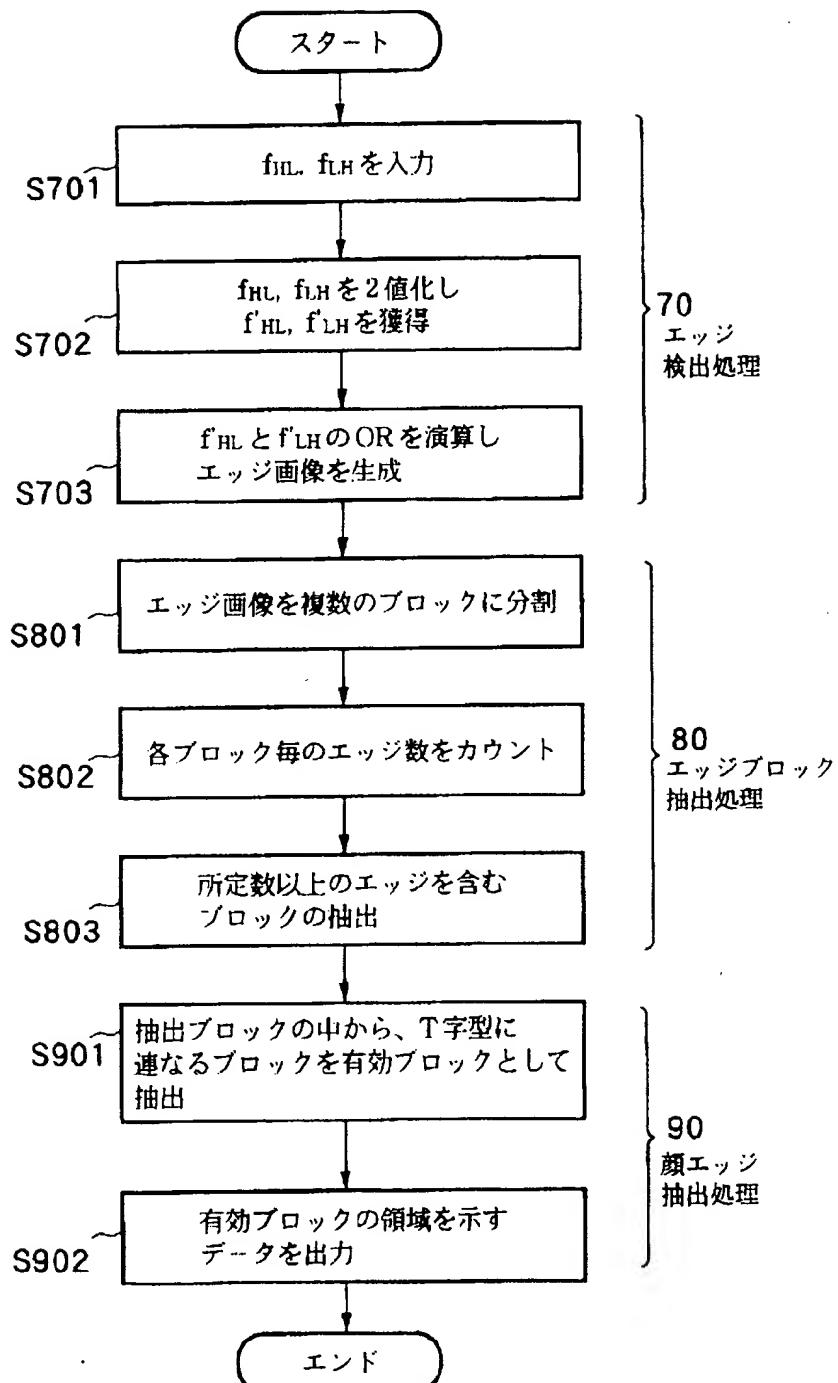


【図16】

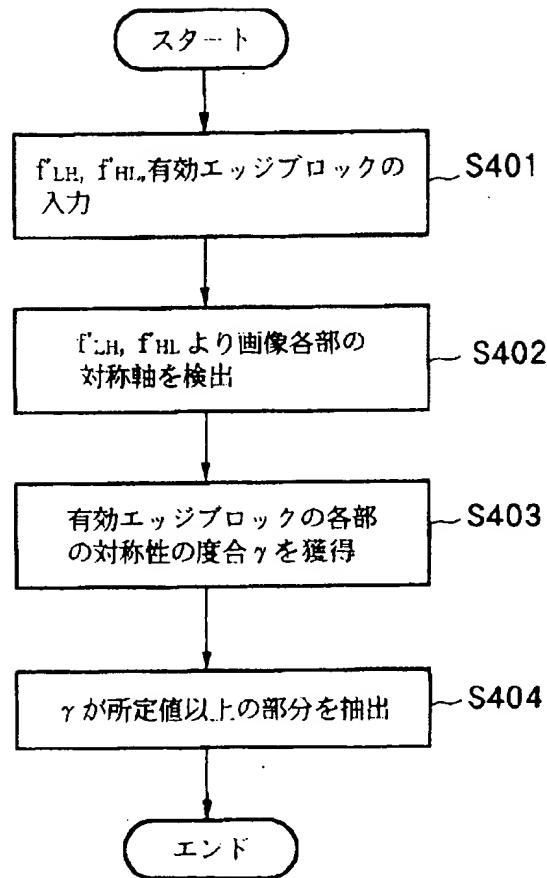


(□ ベストマッチング点)

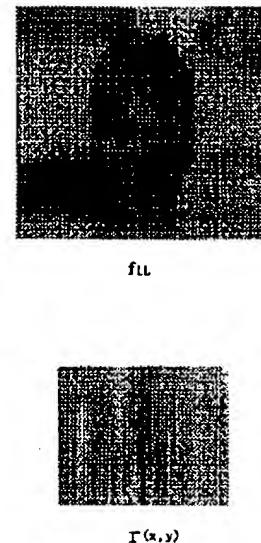
【図6】



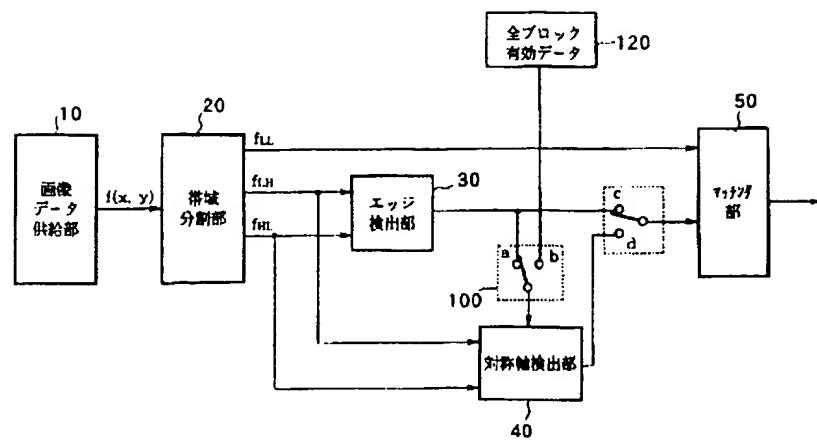
【図11】



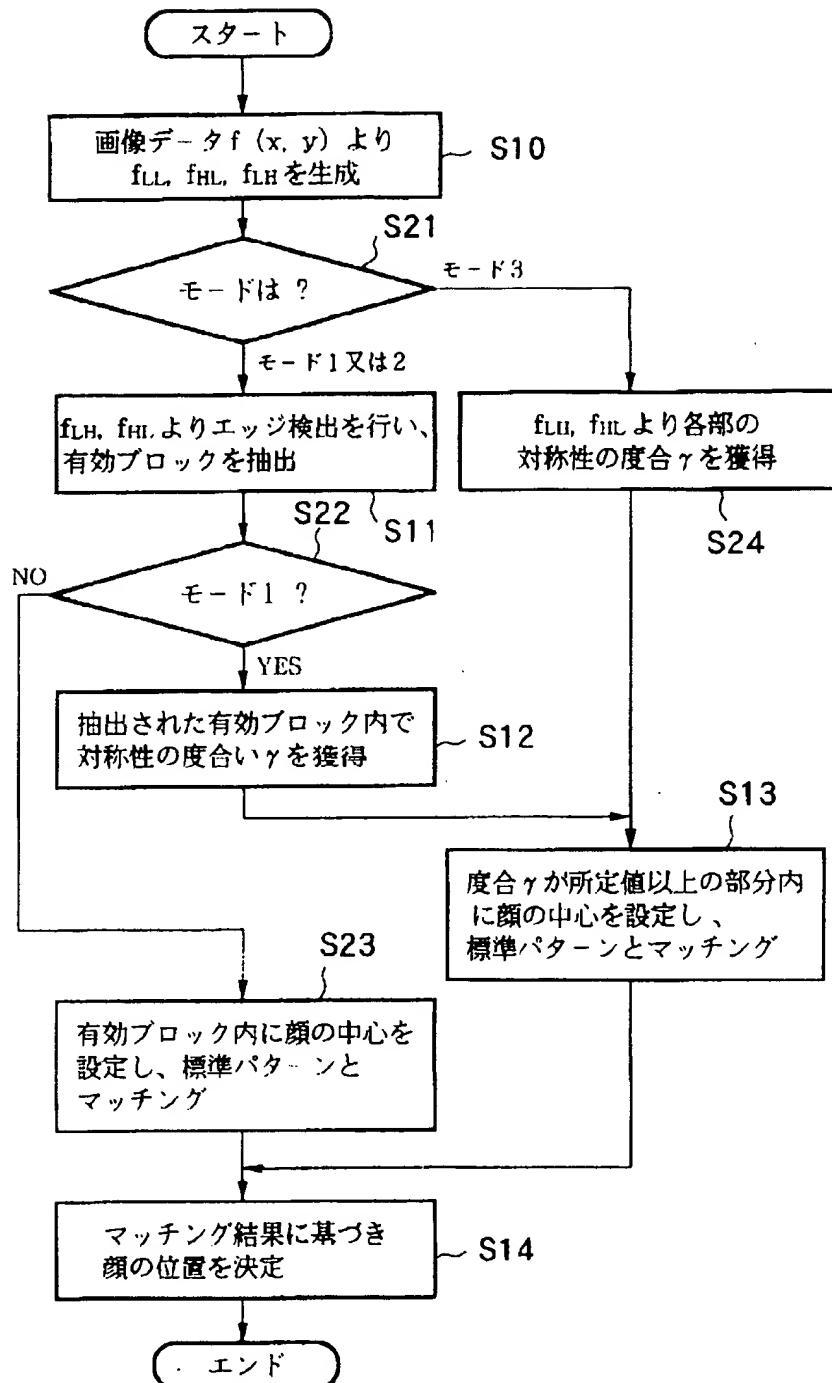
【図12】



【図17】



【図18】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)